

Human sequence of the non-coding RNA gene (including the putative promoter)

1 CTTAGAGTTT CGTGGCTTCA GGGTGGGAGT AGTTGGAGCA TTGGGGATGT
51 TTTTCTTACC GACAAGCACA GTCAGGTTGA AGACCTAACC AGGGCCAGAA
101 GTAGCTTTGC ACTTTTCTAA ACTAGGCTCC TTCAACAAGG CTTGCTGCAG
151 ATACTACTGA CCAGACAAGC TGTTGACCAG GCACCTCCCC TCCCCGCCAA
201 ACCTTTCCCC CATGTGGTCG TTAGAGACAG AGCGACAGAG CAGTTGAGAG
251 GAACTCCCCG TTTTCGGTGC CATCAGTGCC CCGTCTACAG CTCCCCCAGC
301 TCCCCCACC TCCCCCACTC CCAACCACGT TGGGACAGGG AGGTGTGAGG
351 CAGGAGAGAC AGTTGGATTG TTTAGAGAAG ATGGATATGA CCAGTGGCTA
401 TGGCCTGTGC GATCCCAACC GTGGTGGCTC AAGTCTGGCC CCACACCAGC
451 CCCAATCCAA AACTGGCAAG GACGCTTCAC AGGACAGGAA AGTGGCACCT
501 GTCTGCTCCA GCTCTGGCAT GGCTAGGAGG GGGGAGTCCC TTGAACACT
551 GGGTGTAGAC TGGCCTGAAC CACAGGAGAG GATGGCCACG GGTGAGGTGG
601 CATGGTCCAT TCTCAAGGGA CGTCTCCAA CGGGTGGCGC TAGAGGCCAT
651 GGAGGCAGTA GGACAAGGTG CAGGCAGGCT GGCCTGGGGT CAGGCCGGGC
701 AGAGCACAGC GGGGTGAGAG GGATTCTTAA TCACTCAGAG CAGTCTGTGA
751 CTTAGTGGAC AGGGGAGGGG GCAAGGGGG AGGAGAAGAA AATGTTCTTC
801 CAGTTACTTT CCAATCTCTC TTTAGGGACA GCTTAGAATT ATTTGCACTA
851 TTGAGTCTTC ATGTTCCAC TTTAAAACA ACAGATGCTC TGAGAGCAAA
901 CTGGCTTGAA TTGGTGACAT TTAGTCCCTC AAGCCACCAG ATGTGACAGT
951 GTTGAGAACT ACCTGGATTT GTATATATAC CTGCGCTTGT TTTAAAGTGG
1001 GCTCAGCACA TAGGGTCCC ACGAAGCTCC GAAACTCTAA GTGTTTGCTG
1051 CAATTTTATA AGGACTTCCT GATTGGTTTC TCTTCTCCCC TTCCATTTCT
1101 GCCTTTTGTT CATTTCTATC TTTCACTTCT TTCCCTTCCT CCGTCTCTCT
1151 CCTTCTAGT TCATCCCTTC TCTTCCAGGC AGCCGCGGTG CCCAACCACA
1201 CTTGTCGGCT CCAGTCCCCA GAACTCTGCC TGCCCTTTGT CTTCTGCTG
1251 CCAGTACCAG CCCCACCTG TTTTGAGCCC TGAGGAGGCC TTGGGCTCTG
1301 CTGAGTCCAA CCTGGCCTGT CTGTGAAGAG CAAGAGAGCA GCAAGGTCTT
1351 GCTCTCCTAG GTAGCCCCCT CTTCCCTGGT AAGAAAAAGC AAAAGGCATT
1401 TCCCACCCTG AACAACGAGC CTTTTCACCC TTCTACTCTA GAGAAGTGGA
1451 CTGGAGGAGC TGGGCCCCGAT TTGGTAGTTG AGGAAAGCAC AGAGGCCTCC
1501 TGTGGCCTGC CAGTCATCGA GTGGCCCAAC AGGGGCTCCA TGCCAGCCGA
1551 CCTTGACCTC ACTCAGAAGT CCAGAGTCTA GCGTAGTGCA GCAGGGCAGT
1601 AGCGGTACCA ATGCAGAAGT CCCAAGACCC GAGCTGGGAC CAGTACCTGG
1651 GTCCCCAGCC CTTCTCTGTC TCCCCCTTTT CCTCGGAGT TCTTCTTGAA

Fig. 1

1701 TGGCAATGTT TTGCTTTTGC TCGATGCAGA CAGGGGGGCCA GAACACCACA
 1751 CATTTCACTG TCTGTCTGGT CCATAGCTGT GGTGTAGGGG CTTAGAGGCA
 1801 TGGGCTTGCT GTGGGTTTTT AATTGATCAG TTTTCATGTG GGATCCCATC
 1851 TTTTAAACCT CTGTTCAAGG AGTCCTTATC TAGCTGCATA TCTTCATCAT
 1901 ATTGGTATAT CCTTTTCTGT GTTTACAGAG ATGTCTCTTA TATCTAATC
 1951 TGTCCAACTG AGAAGTACCT TATCAAAGTA GCAAATGAGA CAGCAGTCTT
 2001 ATGCTTCCAG AAACACCCAC AGGCATGTCC CATGTGAGCT GCTGCCATGA
 2051 ACTGTCAAGT GTGTGTTGTC TTGTGTATTT CAGTTATTGT CCCTGGCTTC
 2101 CTTACTATGG TGTAATCATG AAGGAGTGAA ACATCATAGA AACTGTCTAG
 2151 CACTTCCTTG CCAGTCTTTA GTGATCAGGA ACCATAGTTG ACAGTTCCAA
 2201 TCAGTAGCTT AAGAAAAAC CGTGTGTC TCTTCTGGAA TGGTTAGAAG
 2251 TGAGGGAGTT TGCCCCGTTT TGTGTTAGTA GTCTCATAGT TGGACTTCTT
 2301 AGCATATATG TGTCCATTTT CTTATGCTGT AAAAGCAAGT CCTGCAACCA
 2351 AACTCCCATC AGCCCCATCC CTGATCCCTG ATCCCTTCCA CCTGCTCTGC
 2401 TGATGACCCC CCCAGCTTCA CTTCTGACTC TTCCCCAGGA AGGGAAGGGG
 2451 GGTGAGAGA GAGGGTGAGT CCTCCAGAAC TCTTCCTCCA AGGACAGAAG
 2501 GCTCCTGCCC CCATAGTGGC CTCGAACTCC TGGCACTACC AAAGGACACT
 2551 TATCCACGAG AGCGCAGCAT CCGACCAGGT TGTCACTGAG AAGATGTTTA
 2601 TTTTGGTCAG TTGGGTTTTT ATGTATTATA CTTAGTCAA TGTAAATGTGG
 2651 CTTCTGGAAT CATGTGTCAG AGCTGCTTCC CCGTCACCTG GCGGTCATCT
 2701 GGTCTGGTA AGAGGAGTGC GTGGCCCACC AGGCCCCCCT GTCACCCATG
 2751 ACAGTTCATT CAGGGCCGAT GGGGCAGTCG TGGTTGGGAA CACAGCATTT
 2801 CAAGCGTCAC TTTATTTTCAT TCGGGCCCCA CCTGCAGCTC CCTCAAAGAG
 2851 GCAGTTGCCC AGCCTCTTTC CCTTCCAGTT TATTCCAGAG CTGCCAGTGG
 2901 GGCCTGAGGC TCCTTAGGGT TTTCTCTCTA TTTCCCCCCTT TCTTCTCAT
 2951 TCCCTCGTCT TTCCCAAAGG CATCACGAGT CAGTCGCCTT TCAGCAGGCA
 3001 GCCTTGCGCG TTTATCGCCC TGGCAGGCAG GGGCCCTGCA GCTCTCATGC
 3051 TGCCCCTGCC TTGGGGTCAG GTTGACAGGA GGTGAGGGG AAAGCCTTAA
 3101 GCTGCAGGAT TCTCACCAGC TGTGTCCGGC CCAGTTTGGG GGTCTGACCT
 3151 CAATTTCAAT TTTGTCTGTA CTTGAACATT ATGAAGATGG GGGCCTCTTT
 3201 CAGTGAATTT GTGAACAGCA GAATTGACCG ACAGCTTTC AGTACCCATG
 3251 GGGCTAGGTC ATTAAGGCCA CATCCACAGT CTCCCCCACC CTTGTTCCAG
 3301 TTGTTAGTTA CTACCTCCTC TCCTGACAA ACTGTATGTC GTCGAGCTCC
 3351 CCCCAGGTCT ACCCCTCCCG GCCCTGCCTG CTGGTGGGCT TGTATAGCC
 3401 AGTGGGATTG CCGSTCTTGA CAGCTCAGTG AGCTGGAGAT ACTTGGTCAC

Fig. 1 (cont'd 1)

3/21

3451 AGCCAGGCGC TAGCACAGCT CCCTTCTGTT GATGCTGTAT TCCCATATCA
 3501 AAAGGCACAG GGGACACCCA GAAACGCCAC ATCCCCCAAT CCATCAGTGC
 3551 CAAACTAGCC AACGGCCCCA GCTTCTCAGC TCGCTGGATG GCGGAAGCTG
 3601 CTACTCGTGA GCGCCAGTGC GGGTGCAGAC AATCTTCTGT TGGGTGGCAT
 3651 CATTCAGGC CCGAAGCATG AACACTGCAC CTGGGACAGG GAGCAGCCCC
 3701 AAATTGTCAC CTGCTTCTCT GCCCAGCTTT TCATTGCTGT GACAGTGATG
 3751 GCGAAAGAGG GTAATAACCA GACACAAACT GCCAAGTTGG GTGGAGAAAG
 3801 GAGTTTCTTT AGCTGACAGA ATCTCTGAAT TTTAAATCAC TTAGTAAGCG
 3851 GCTCAAGCCC AGGAGGGAGC AGAGGGATAC GAGCGGAGTC CCCTGCGCGG
 3901 GACCATCTGG AATTGGTTTA GCCCAAGTGG AGCCTGACAG CCAGAACTCT
 3951 GTGTCCCCCG TCTAACCACA GCTCCTTTTC CAGAGCATTC CAGTCAGGCT
 4001 CTCTGGGCTG ACTGGGGCAG GGGAGGTTAC AGGTACCAGT TCTTTAAGAA
 4051 GATCTTTGGG CATATACATT TTTAGCCTGT GTCATTGCCC CAAATGGATT
 4101 CCTGTTTCAA GTTCACACCT GCAGATTCTA GGACCTGTGT CCTAGACTTC
 4151 AGGGAGTCAG CTGTTTCTAG AGTTCCTACC ATGGAGTGGG TCTGGAGGAC
 4201 CTGCCCCGTG GGGGGGCAGA GCCCTGCTCC CTCCGGGTCT TCCTACTCTT
 4251 CTCTCTGCTC TGACGGGATT TGTGATTCT CTCCATTTTG GTGTCTTTCT
 4301 CTTTTAGATA TTGTATCAAT CTTTAGAATA GGCATAGTCT ACTTGTTATA
 4351 AATCGTTAGG ATACTGCCTC CCCCAGGGTC TAAATTTACA TATTAGAGGG
 4401 GAAAAGCTGA ACACTGAAGT CAGTTCTCAA CAATTTAGAA GGAAAACCTA
 4451 GAAAACATTT GGCAGAAAT TACATTTTGA TGTTTTTGAA TGAATACAG
 4501 CAAGCTTTTA CAACAGTGCT GATCTAAAAA TACTTAGCAC TTGGCCTGAG
 4551 ATGCCTGGTG AGCATTACAG GCAAGGGGAA TCTGGAGGTA GCCGACCTGA
 4601 GGACATGGCT TCTGAACCTG TCTTTTGGGA GTGGTATGGA AGGTGGAGCG
 4651 TTCACCAGTG ACCTGGGAGG CCCAGCACCA CCCTCCTTCC CACTCTTCTC
 4701 ATCTTGACAG AGCCTGCCCC AGCGCTGACG TGTCAGGAAA ACACCCAGGG
 4751 AACTAGGAAG GCACTTCTGC CTGAGGGGCA GCCTGCCTTG CCCACTCCTG
 4801 CTCTGCTCGC CTCGGATCAG CTGAGCCTTC TGAGCTGGCC TCTCACTGCC
 4851 TCCCCAAGGC CCCCTGCCTG CCCTGTCAGG AGGCAGAAGG AAGCAGGTGT
 4901 GAGGGCAGTG CAAGGAGGGA GCACAACCCC CAGTCCCCGC TCCGGGCTCC
 4951 GACTTGTCAG CAGGCAGAGC CCAGACCCTG GAGGAAATCC TACCTTTGAA
 5001 TTCAAGAACA TTTGGGGAAT TTGGAAATCT CTTTGCCCCC AAACCCCAT
 5051 TCTGTCTTAC CTTTAATCAG GTCCTGCTCA GCAGTGAGAG CAGATGAGGT
 5101 GAAAAGGCCA AGAGTTTGG CTCCTGCCCC CTGATAGCCC CTCTCCCCGC
 5151 AGTGTGTTGTG TGTCAAGTGG CAAAGCTGTT CTTCTGGTG ACCCTGATTA
 5201 TATCCAGTAA CACATAGACT GTGCGCATAG GCCTGCTTTG TCTCCTCTAT

Fig. 1 (cont'd 2)

4/21

5251 CCTGGGCTTT TGTTTTGCTT TTTAGTTTTG CTTTGTAGTTT TTCTGTCCCT
 5301 TTTATTTAAC GCACCGACTA GACACACAAA GCAGTTGAAT TTTTATATAT
 5351 ATATCTGTAT ATTGCACAAT TATAAACTCA TTTTGCTTGT GGCTCCACAC
 5401 ACACAAAAAA AGACCTGTTA AAATTATACC TGTGTGCTTAA TTACAATATT
 5451 TCTGATAACC ATAGCATAGG ACAAGGGAAA ATAAAAAAG AAAAAAAGA
 5501 AAAAAAACG ACAAATCTGT CTGCTGGTCA CTTCTTCTGT CCAAGCAGAT
 5551 TCCTGTCTTT TTCTCGCTT CTTTCAAGGG CTTTCTGTG CCAGGTGAAG
 5601 GAGGCTCCAG GCAGCACCCA GGTTTTGCAC TCTTGTCTT CCCGTGCTTG
 5651 TGAAAGAGGT CCCAAGGTTT TGGGTGCAGG AGCGCTCCCT TGACCTGCTG
 5701 AAGTCCGGAA CGTAGTCGGC ACAGCCTGGT CGCCTTCCAC CTCTGGGAGC
 5751 TGGAGTCCAC TGGGGTGGCC TGAATCCCCC AGTCCCCTTC CCGTGACCTG
 5801 GTCAGGGTGA GCCCATGTGG AGTCAGCCTC GCAGGCCTCC CTGCCAGTAG
 5851 GGTCCGAGTG TGTTCATCC TTCCCACTCT GTCGAGCCTG GGGGCTGGAG
 5901 CGGAGACGGG AGGCCTGGCC TGTCTCGGAA CCTGTGAGCT GCACCAGGTA
 5951 GAACGCCAGG GACCCAGAA TCATGTGCGT CAGTCCAGG GGTCCCCTCC
 6001 AGGAGTAGTG AAGACTCCAG AAATGTCCCT TTCTTCTCCC CCATCCTACG
 6051 AGTAATTGCA TTTGCTTTTG TAATCTTAA TGAGCAATAT CTGCTAGAGA
 6101 GTTTAGCTGT AACAGTTCTT TTTGATCATC TTTTCTTAA AATTAGAAAC
 6151 ACCAAAAAAA TCCAGAACT TGTCTTCCA AAGCAGAGAG CATTATAATC
 6201 ACCAGGGCCA AAAGCTTCCC TCCCTGCTGT CATTGCTTCT TCTGAGGCCT
 6251 GAATCCAAA GAAAAACAGC CATAGGCCCT TTCAGTGGCC GGGCTACCCG
 6301 TGAGCCCTTC GGAGGACCAG GGCTGGGGCA GCCTCTGGGC CCACATCCGG
 6351 GGCCAGCTCC GCGGTGTGTT CAGTGTTAGC AGTGGGTCAT GATGCTCTTT
 6401 CCCACCCAGC CTGGGATAGG GGCAGAGGAG GCGAGGAGG CGTTGCCGCT
 6451 GATGTTTGGC CGTGAACAGG TGGGTGTCTG CGTGGCTCCA CGTGGCTGTT
 6501 TTCTGACTGA CATGAATCG ACGCCCGAGT TAGCCTCACC CGGTGACCTC
 6551 TAGCCCTGCC CGGATGGAGC GGGGCCACC CGGTTCACTG TTTCTGGGGA
 6601 GCTGGACAGT GGAGTGCAA AGGCTTGCAG AACTTGAGC CTGCTCCTTC
 6651 CCTTGCTACC ACGGCCTCCT TTCCGTTTGA TTTGTCACTG CTTCAATCAA
 6701 TAACAGCCGC TCCAGAGTCA GTAGTCAATG AATATATGAC CAAATATCAC
 6751 CAGGACTGTT ACTCAATGTG TGCCGAGCCC TTGCCCATGC TGGGCTCCCG
 6801 TGTATCTGGA CACTGTAACG TGTGCTGTGT TTGCTCCCTT TCCCCTTCCT
 6851 TCTTTGCCCT TTACTTGTCT TTCTGGGGTT TTTCTGTTG GGTTTGGTTT
 6901 GGTTTTTATT TCTCCTTTTG TGTTCCAAAC ATGAGGTTCT CTCTACTGGT
 6951 CCTCTTAACT GTGGTGTGA GGCTTATATT TGTGTAATTT TTGGTGGGTG

Fig. 1 (cont'd 3)

7001 AAAGGAATTT TGCTAAGTAA ATCTCTTCTG TGTTTGAAGT GAAGTCTGTA
7051 TTGTAAGTAT GTTTAAGTA ATTGTTCCAG AGACAAATAT TTCTAGACAC
7101 TTTTCTTTTA CAAACAAAAG CATTCGGAGG GAGGGGGATG GTGACTGAGA
7151 TGAGAGGGGA GAGCTGAACA GATGACCCCT GCCCAGATCA GCCAGAAGCC
7201 ACCCAAAGCA GTGGAGCCCA GGAGTCCCAC TCCAAGCCAG CAAGCCGAAT
7251 AGCTGATGTG TTGCCACTTT CCAAGTCACT GCAAAACCAG GTTTTGTTC
7301 GCCCAGTGGA TTCTTGTTTT GCTTCCCTC CCCCCGAGAT TATTACCACC
7351 ATCCCGTGCT TTAAAGGAAA GGCAAGATTG ATGTTTCTT GAGGGGAGCC
7401 AGGAGGGGAT GTGTGTGTGC AGAGCTGAAG AGCTGGGGAG AATGGGGCTG
7451 GGCCCAACCA AGCAGGAGGC TGGGACGCTC TGCTGTGGGC ACAGGTCAGG
7501 CTAATGTTGG CAGATGCAGC TCTTCCTGGA CAGGCCAGGT GGTGGGCATT
7551 CTCTCTCCAA GGTGTGCCCC GTGGGCATTA CTGTTTAGA CACTTCCGTC
7601 ACATCCCACC CCATCCTCCA GGGCTCAACA CTGTGACATC TCTATTCCCC
7651 ACCCTCCCCCT TCCCAGGGCA ATAAATGAC CATGGAGGGG GCTTGCACTC
7701 TCTTGCGTGT CACCCGATCG CCAGCAAAAC TTAGATGTGA GAAAACCCCT
7751 TCCCATTCOA TGCGGAAAC ATCTCCTTAG AAAAGCCATT ACCCTCATTA
7801 GGCATGGTTT TGGGCTCCCA AAACACCTGA CAGCCCCCTC CTCCTCTGAG
7851 AGGCGGAGAG TGCTGACTGT AGTGACCATT GCATGCCGGG TGCAGCATCT
7901 GGAAGAGCTA GGCAGGGTGT CTGCCCCCTC CTGAGTTGAA GTCATGCTCC
7951 CCTGTGCCAG CCCAGAGGCC GAGAGCTATG GACAGCATTG CCAGTAACAC
8001 AGGCCACCCCT GTGCAGAGG GAGCTGGCTC CAGCCTGGAA ACCTGTCTGA
8051 GGTGAGGAGA GGTGCACTTG GGGCACAGG AGAGGCCGGG ACACACTTAG
8101 CTGAGATGT CTCTAAAGC CCTGTATCGT ATTCACCTC AGTTTTTGTG
8151 TTTTGGGACA ATTACTTTAG AAAATAAGTA GGTGGTTTTA AAAACAAAA
8201 TTATTGATTG CTTTTTTGTA GTGTTCAAG AAAAGGTTCT TTGTGTATAG
8251 CCATATGACT GAAAGCACTG ATATATTTAA AAACAAAAG CAATTTATTA
8301 AGGAAATTTG TACCATTTCA GTAAACCTGT CTGAATGTAC CTGTATACGT
8351 TTCAAAAACA CCCCCCCCCC ACTGAATCCC TGTAACCTAT TTATTATATA
8401 AAGAGTTTGC CTTATAAATT TA

Fig. 1 (cont'd 4)

6/21

Murine sequence of the non-coding RNA gene (including the putative promoter)

```

1  CTTAGAGTTT CGTGGCTTCG GGGTGGGAGT AGTTGGAGCA TTGGGATGTT
51  TTTCTTACCG ACAAGCACAG TCAGGTTGAA GACCTAACCA GGGCCAGAAG
101 TAGCTTTGCA CTTTTCTAAA CTAGGCTCCT TCAACAAGGC TTGCTGCAGA
151 TACTACTGAC CAGACAAGCT GTTGACCAGG CACTCCCCCC AACAAATATCC
201 TCCCTCTTCC CCCCCCCCAC CCCC GCCCCG TGTGCTCGTT AGGGCAATTG
251 AAAGGACACT CCCATTTTTG GTGCCATTGA TGCCCTGTCC ATAATAGCTT
301 CCCTGACTTT TACACCACCC CAACTCCCAA TCTGAAGGAC TGGGAGGTGT
351 GATGCAGGAG AACTATGGG ACTCTTGGGA GAAGACTATG GAGTTGGCCA
401 GTGATTAAAG CCCACTAATT CCAACTGTGG TAGCACAGAT CTGGCTCCAC
451 ATCAACCCAA TCCAAAAGTG ACAAGGATAT TTTGCAAAAA AAGAAAGTGG
501 CACCTGTCTG ATCCAGCTCT GACATGGCTA GAGGTGAGTC CTAAACTGAT
551 GGCTTATAAA CTAGCCTGAG CCACAGAAGA GTATGGCCCA GAGTGAAGTG
601 TCATCATCTG TTCACAAGGC ATGCTCCCTT AGAAGATAAT GCTAAAGAGG
651 TGCCATGGAG GCAGCAGGAC AAGTACAGG CAGGCTAGGT GGAGTCAAGC
701 CAGGCCTAGT GCCACAGAAC AAGAGAGCAG TCTGACTAGT AATTAAGAGG
751 GAAGAAAGGA AATATCTCTT CCAATTACTT TCCAGTTCTC CTTTAGGGAC
801 AGCTTAGAAT TATTTGCACT ATTGAGTCTT CATGTTCCCA CTTCAAAACA
851 AACAGATGCT CTGAAAGCAA ACTGGCTTGA AATGGTGACA CTGTCCACAA
901 AGCCACCAGA CATGGCAGTG TTCAGAACTA CCTGTATCTG TATATACCTG
951 CGCTTGTTTT AAAGTGGGCT CAGCACATAG GATTCCCAAG AAGCTCCGAA
1001 ACTCTAAGTG TTTGCTGCAA TTTTATAAGG ACTTCCTGAT TGCTTTCTCT
1051 CTCGTCTTTC CATTTCTTCC TTCCTTCCAT TTCATGCTTT CATTTCTTCC
1101 CCTAGCTTCT AGTTGTTTCT TCTGTTCCAG GCAGCTGCAG TGCTGAACCA
1151 CATGGTTACC TAACAGCAGT CAGCTGCAGC CCTAGGATTC TTCCTGCCCT
1201 TTAAC TTCCC ATTGCCAGTG CCAGGTATCA TATTTAACCT TGAGCAAGAG
1251 CTGGGCTCTT TTGAGCCCTC CCTAACCTCT GTGAAGAAGA ACAAGAAGGT
1301 AGGAAGCTCT TGCTCTTGCT AAGAAAAATG TCAAAAGGCT TTCAGACCTT
1351 AAACAATGAG CCTTTTCACC TTTTACTCTA GAAAAGTGGA CTAGAAAATC
1401 TGGGTCACAT TGGGTAGCTG AAGGAGATAC AGAGGCCCTT ATGGCCTGCC
1451 AGAGTCGTTG CATGGCCCAA CAGGGGCTCC ATGCCCACTA CCCTTGACCC
1501 TACTCAGAAA TCTAATGTCA TACTTAGTGT GGGCAGGGGA CCTGTCAGGA
1551 CAGATGCAGA CCTAAGCAGG GAGTGACACC AGGGCCCTTG GCCCTTCTTC
1601 TGACAAACAT ACACATCCCA AGTCTTTTTT TAGTGGAATT CTTAACCTCT
1651 TGCTCACTGG GGACTGGGAA GCATCAGCAC ATCCCATATT TCAAACCTCTG

```

Fig. 2

7/21

1701 CTCCATAAGT ACAGTGGTGA ATTTTATAGA CTTGACTTTG CTGTGGGGTT
 1751 TTAATTGGTC AGTTTAAATT TGGGATCCCA AAGTTTAAAC CTCCATTTCAG
 1801 GAAGTCCTTA TCTAGCTGCA TATCTTCATC ATATTGGTAT ATCCTTTTCT
 1851 GTGTTTACAG AGATGTCTCA TATCTATCGA AATCTGTCTG AGAAGTACCT
 1901 TATCAAAGTA GCAAAAGAGA CAGCAGTCTT ATGCTTCCAG AAACACCCAC
 1951 AGGCACGTCC CATGTGAGCT GCTGCCATGA ACTGTGAGT GTGTATTGTC
 2001 TTGTGTATTT TCGTTAACGT TCCCCAGCTT CCTTCCTGCG GTGTAATCAT
 2051 GGAAGAGTGA AACATCATAG AAATCGTCTA GCACCTCCTG GCCAGTCCCT
 2101 AGTGATCAGG AACCGTAGTT GACAGTTCCA ATTGATAGCT TAAGATAAAA
 2151 CCATGTTTGT CTCTTATGGA ATGGTTAGAA CTAAGTGAGA GATCTTGCCC
 2201 CATTCGTGTTT GCCGATTCAT AGTTGGACTT TTAGTGTATT TGTATCCATT
 2251 TCCTTGCTGCT ATAAAGCAA ACCCTGCAAC CAGCTTTCTG TCAGGCAGTC
 2301 CTTTTGCGCTG CTCTGCTTTT GATCCTCTTA GTCTTGCTTC TGGTTCCCTCC
 2351 CTGGAGAGGG AGGAGGGGTC AGAAGAGGAA TTCTGGAGGA TCCAGGATAT
 2401 GTCCTTCTGA ACTCCTGCTT CTTCCAGTGA CAAAGGCCC CTAAGTCCCC
 2451 ACCCCAACTT GCCCATGCA CTCCTCTAGG ACACCTTCC ATACTTTTCA
 2501 CAACACCTAG CCAGGTTGAC ACCAAGTTGT TTATTGTTGGT CTGCTTGGAA
 2551 TTTTACCTGT TAGGCTTACT TAGTCCAATC AAATGGACTC CAAGTTGGGT
 2601 ATCCCTCATC TTTGGAGAC AACCTAGGCT GATTAGATAT TTACTTTTGG
 2651 GATTGCAGCA CTTTGGGTGC CTTTTTCTT TTACTTGGGT TTTATCTGCA
 2701 GCTCCCTCAC CACCACCACC ACCCCCCACT TACCTGTATG TAGAACTGAT
 2751 TTCAAACTG CAGGTGGTGG TAACTGCAGC TTCTTAGGGT TTTCTTCACT
 2801 TCTTGCTTCT TTCCCCATTC CCTCATCCAC AAATAAGGGC ATCACAAGTC
 2851 AGTCTCCTTT AAGCAGGCAG CTTTGGTGGG GTTTTCCCC TGGAAAGCCAG
 2901 GGACCCTGTC AGGCTGCCTC TGCCTTGTGG TCAGGTTGAC AGGAGGTTGG
 2951 AGGGAAAAGC CTTAAGTCAT GGGATTCTCA CCAGCTGTGT CTGGCTCAGA
 3001 CCTGGAATGT GACCTTTATT TTGTTGTATT TGAACATTGT AAAGTGTGGG
 3051 TGGTACCTTA AACTGAATAT GTGAAGAATC CAGAACTGA CCAACAGCTT
 3101 TCAGATACCT GGGGCTAGGT CACTAAGGTC ACATCCAGTC TTCCCTACCC
 3151 TGTCTAGTT GTTAGCTACT ACCTCTCCCA GATAGATTGC TGTATATCCT
 3201 CCAACTATGA TCATCCTGGC CCAAGCTTGC CTGTTCTTGA GTCTGTCTTA
 3251 ACCAGTGGAA CTGCTGCCCT TGGTGTGCAG TGAGTTGAGG ACTCTTGGTC
 3301 ACAGCCAGGC TCTAGTAGTA CAGCTCCTTT CTGCTGGTGC TGTATTTCCA
 3351 TATCAAAAGG CACAGGGGAG ATCTAGAAAT GCCATCTCCC CCAGTCCATC
 3401 AGTGCCAAAC AAGCCCATGA TCCCAGCATG GGTACAGACA ACTCTGTTCA

Fig. 2 (cont'd 1)

3451 GTGCTATCAC AACAGACTAG AGGCCATGAA CATTGGACGT GGGAAACCAGA
3501 GCAACCCGAA TTGCTGCTGC TTTATTTCAGC TTTCCGTTGC TCTGACAATG
3551 ATAAACAAG GCAGTAACTT AAAACAGACT GCCAGGTTTG GCAGAGAAAG
3601 GAAATTCCCTT AGCTGACAGC ACCTCTGGAT TTTAAATAGG TTGTAATAAG
3651 TGGCTCAAAC CCATCCAGGA AAAAGCAAAA GGGTTAGAAC TGACCAGATG
3701 AGACCAGCCT GATTTTCATGC AGCCCCAAATG GAGTCCAGCT GTCTGAACTC
3751 TGCAGCACTT CTCTACTACA GTCTCCTAGA GCATTCCAGC CAGGCTCTTC
3801 AGGCTGAGGA GACATCACAG GTGCCAGTTC TTCAGAGAGA CTTTTGTGCA
3851 TCAGTTCATA GCCTATATCT TTGCCCAAGA TTGTAGATTTC AGGTTAACAC
3901 TACAGATTCT AGGGCAGATG ACTGAGACTC AGAAAAAAG CCCCTGTGGA
3951 CTGTGGTATA GCGAAGTACA AAAACTGAAG GGGGCTAGGG CAGATGCCGC
4001 ATGCCTCATG CCAGAGCCAA GCCCTCTGCT CCATCCACAT CCTTTTCTGG
4051 CTCCTTCTTC CTGCTCTCTG CTTCACTGAA CCAGCCCCAC TCTGAAGAGA
4101 TTTGTTGATT CTCTCCATTT TTATGTCTTT CTCTTTTTAGG TACTATATAG
4151 AAAAGGCTTA GTCTAATTGT TATAAATTGC TAGAATACTG CCTCCCCCAG
4201 GGTCTAAAAA TATATGCTAA AGGGGAAAPC TTGAACACTG AAACCAGTTC
4251 TCAACAATTT AGAAGGAAAA CCTTGAAAPC ATTTAACAPA AAATTATATT
4301 TTAATGTTTA TGAATAAGAG GAGGCTTTTG AAAAAATGTT GATCTATAAA
4351 TACTTACTTT AGGCCTGAGG TGTCTAATGA GTGAACTGAG CAATGGGAAC
4401 TCAAGGCTGA AGCCTCCTGC ATCAGAGGAG GTAGAACCAG GAGCCTCTTG
4451 AGATTTGAGG TGTTTTAGCA TTGGAAAGCC ACTCTTTGGG TAGCTGGCCC
4501 CAGAAACTAC TTCTGACCTT GTCATTTGGA ATGGAGGTTA GTGGTCTGCC
4551 AGATGCCAAA GCTGCACTGAG ACCAGCTCTT GGTTTATCAA TTTGAACACT
4601 CAGTAACCTA GAAGGCCCAG CACAAAGTGT CTGCTCTCTT CTTAACTGAG
4651 CCTGCCCCAG CACTACTGCA CAAATTAGGG AGGGTCTACT TCCTACAGAG
4701 CATCCCTCCC TGGGCCCCCT CCCATCCTTT GACTCTACC TACCTGACCT
4751 TCAGGATCTT GGCACATACG AATGGCTGT GTAGCAAGCA CTTTGGCATG
4801 CCGTCCTAAA CTTACCCCAG AGCCTCTCCC TGCCTCCTTA AGCCAGTCTG
4851 CCTGTCTTCT GGGGAGGTGT TAGAGCCCAT AGAATGGAGA GGAGAAAGAA
4901 AAGAGGAAGA GGCAGGCAGG TAGTAAAAAG GCTCTGGGAG GAAAGACAGC
4951 CTCCTAGGCT TTGCACAAGC AGGACTCAGC CCCTTGTGGG AACTAAGTGC
5001 CATCTTGGAG TTTAAGAACA TTTGGACAAG TTGCAAATGA CCTTTGCTCC
5051 TTGCTCCTCT CACCTTTTAT GGGGCCCTGC TTAGCACTGA AAGCAAATGC
5101 GCTGAAAAGG CAAAGAGGTT TGGCTCCTGC CCACTGATAG TCCTTTCCCT
5151 GCAGTGTTTG TGTGTCAAGT GGCAAAGCTG TTCTTCCTGG TGA CTCTGAT
5201 TAGATCCAGT AACTTAAGAG ATTTGTATGC ATAGGTCTGC TTTGACTCTT

Fig. 2 (cont'd 2)

9/21

5251 CTATTCTGGG CTTTGTGATT GTTTTTCAGT TTTGCTTTTA GTTTTCCTAT
 5301 TTTTATTTTA TGCACCAACT AGACACACAA AGCAGTTGAA TTTATATATA
 5351 TATATATATA TATATATCTG TATATTTTAC AATTATAAAC TCATTTTGCT
 5401 TGTGACGCCA CACACACACA AAAAGAAAAA CCTTTTAAAA TTATACCTGT
 5451 TGCTTAATTA CAATATTTCT GATAACCATA GAGTAGGACA AGGGAAAAAA
 5501 TTTAAAAAAA AAAAAA AAAAGAAAAA ACATCTGTCT GCTGGTCACT
 5551 TCTTCAATCC AAGCAGATCT GTGATCTTTC CTCGCGTCTT TCAAAGACTT
 5601 CCCTGTGCTA AGTGAAGGAA GCTCCAGGCT GCACCCAGGT TTTGTGCTTT
 5651 GTTTCCTCCTC TGTGTGAAA GGGGCCCCAA GATTCTGGGT ACAGGACAGT
 5701 TCATTTTCAGC ATGGGGTCAG GAGACAAGAG CACTCCCTTT ACATGCTGAC
 5751 GTACAGAACT TAGTGGGAAT AGCCTAGTCC CCACCTCTAG GGATGGGGAG
 5801 CTAGCATGCA TGGGGGTGAC CCAACTCCCT CCACCTTTCC CTGGCCAGGA
 5851 AGAGCCTGTG TACAGTAAGT CTGACAAGCT TTCCCCAGTT AGCAGGGCTC
 5901 AGAGCATTTA AAAACCCCTC AAACCTTTGCT GAGTCTAGGG ACTAGAGAGA
 5951 AGATAGAAGA TTTGGTCTAT CTCCAAGGTG TGTAAGCTGT ACCAGGTAGA
 6001 ATGCCAGGGA CCCCAGAACC ACATCCAACA GCCCAATGGG TCTCCTCCAG
 6051 AAAGTAGTGA AGACTCCAGA AACATCCCTT TCTCTTCTCC CTGCTCCCAT
 6101 GAGTAACTGC ATTTGCTTTT GTAATCCTTA ATGAGCATTA TCTGCTAAAA
 6151 AAAAAAATT AGCTGTAACA GTTCTTTTGT CAAAAGGATC ATTCTTAAAT
 6201 AATTAAAAAC ACCCCCCCCC CAAAAAAG TCCAGAACCT TGTCTTCCA
 6251 AAGCAGAGAG CATTATAATC AGGGCCAAA TCTGTCCAC ACCTCTACCC
 6301 CATCTCCTCA TGATTGCTGC TTCTAAGGCC AGAATACAGC AAAGATATTT
 6351 GTAGGCCCTT TGGGTGACTG GGCTACCCTT GGAGCTCTTG GAAGATGGGC
 6401 TGGGAAGCC TCTGAGACCC TATCCTAGGG CCTTGCTCTA GGGAGTAATC
 6451 AGTATTAGTA GAGTGTACA ACATTATTCC CCAGCCGGCA TGAGATGGGG
 6501 GCAGAAGAAG CCAAGGGTT GTCTCCACTG CTACTTACTT GGCCACTGAC
 6551 AGGTAGGTGA CCATGTATGT CCATATGCAT GTTTTATGGC TGATGTGAGA
 6601 TCAGCACCCA AGTTAGCTTC ACCTGGTGAC CTCTAACCTT GCCTGGATGG
 6651 AGCAGGCCAC CTGGTTCAAT GTTCTTGGGC AGCTGGACAA TGGAGTGCAA
 6701 AAGGCTTACA GAACCTGAAG CCTTTTCCTT ACTTTGCTAG CACGGCCTCC
 6751 TTTTCCATTT GATTTGTCAC TGCTTCAGTC AATAACAGCC GCTCCAGAGT
 6801 CAGTAGTTGA TGAATATATG ACCAAATATC ACCAGGACTG TTA CTCAACG
 6851 TGTGCCGAGC CCTTTCCTTG TGCTGGGCTC CCTGTGTACC TGGACACTGT
 6901 AATGTGTGCT GTGTTTGCTC TCCTTCCTCT TCCTTCCTTG CCCTTTCCTT
 6951 GTCTTCTGG GGTTTTCTG TTGGGTTTG TTTGGTTTAA TTTTTCCTTT

Fig. 2 (cont'd 3)

10/21

7001 TGTGTTCCAA ACATGAGGTT TTCTCTACTG GTCCTCTTTA ACTGTGGTGT
7051 TGAGGCTTCT ATTTGTGTAA TTTTGGGTGG GTGAAAGGAA CTTTGCTAAG
7101 TAAATCTCTT CTGTGTTTGA AATGAAGTCT GTATTGTAAC TATGTTTAAA
7151 GTAATTGTTT CAGAGACAAA TGCTTCTAGG TACATTTTCA TTACAAACAA
7201 AGCATTGTA GGGAGGGAAG TGGTGAATAA GACAAGAGGG GCAATCTGAA
7251 TTGATCCCTG CCCAGATCAG CCAGAAGCTA CCAAAGTTA AGCACTGGTT
7301 TTCCATTCCA AGTCAAGAGA CTGAAGCTGA TGTTTTGCCA TTTTCAAAGT
7351 CAAAGCAAAA CCAGCTTTTC CACCCAATGG ATTCTTTGCT TCTCCTTCCC
7401 AGATTATTAC TACTGCTGTA ATAATCTAGG AGTGCCAGGA GGGAAAGGAG
7451 TATTAACACA GAGCTGTGCT CACTGAGTAT GGAAAGGCTT GGTCTGAGTT
7501 TTCAGGAGGA TGACCCACTG TGGACATGGG GAGAAGACAG AAGATAAATT
7551 AGCCGCTCCC TGCCTAAGAT ACCTCTTAAT AGATAAGTCA AGGCCATGGA
7601 CATTATTGTC TACAAGGCAT GTTTCAAAGA CATGACCAGT CAGGACACTT
7651 CTGTCATACT CCATGTTGCC CCTAGTACA CAGTACTAAT CTGATATCTC
7701 TGTTCCTGCC ATGCCTGGGG GATAAATGA TAGCAGAGAC TCCTTTCTTT
7751 CAATGTGATC TAATTCCCAA CAAAATCTGG GCCTGAGATA CCACCTGTTT
7801 CTATGGCAAA CATCCTCAGT AAGTGTTAT TCTCATTGCA GATTGTTCCA
7851 GCCTAATGTA AGAGGAACAG AGCAGTGTTT CCTTGGAGCC TCATGTGGAC
7901 AGTTCTACCT GTAGTGACCA GTTGGCTATA GTAGTTATTA GCTGGAACAA
7951 CCAGACAGGG TACATGCCCC CTCCAAATC CATGTTGTAC TCCCCTCTGC
8001 CAGCCAGGGG GGGTGAGATC TGTAGAATAG TGCAGCCAGT GACAAGCCAC
8051 CTTGTGTTTG TCACCAGCTC AAAAATCAT CTAAGGTTGG GAGCAGGCAG
8101 ACAAGGCAGA GAGAAAGATC CAGGACAGAC CTAGCTGGGC TGGAGGGGTC
8151 TTGAAAAGCC CTCTGTCGTA TTCACCTTCA GTTTTTGTGC TTTGGGACAA
8201 TTACTTTAGA AAATAAGTAG GTCGTTTTAA AAACAAATA TTGATTGCTT
8251 TTTTGTAGTG TTCAAACAA AAGGTCTTTT GTGTATAGCC AAATGACTGA
8301 AAGCACTGAT ATATTTAAAA ACAAAGGCA ATTTATTAAG GAAATTTGTA
8351 CCATTTTCAGT AAACCTGTCT GAATGTACCT GTATACGTTT CAAAACACA
8401 CCCCCTGAA CCCCTGTAAC CTATTTATTA TATAAAGAGT TTGCCTTATA
8451 AATTTACATA AAAA

Fig. 2 (cont'd 4)

Human
Mous

1 CTAGAGATTTCGTGGCTTCAGGGTGGGAGTAGTTGGAGCATTTGGGATGT
1 -----G-----
51 TTTTCTTACCGACAAGCAGCTCAGGTTGAAGACCTAACAGGGCCAGAA
50 -----
101 GTAGCTTTGCACCTTTCTAACTAGGCTCCTTCAACAGGCTTGCTGCAG
100 -----
151 ATACTACTGACCAGACAAGCTGTGACCAGGCACCTCCCC.....
150 -----TC-----CAACAATATC
191TCCCGCCCAACCTTTCCCCCATGTGGTCTGTAGAGACAGA
200 CTCCTCTCTC-----C---CC---CCCG---G---C-----
232 GCGACAGAGCAGTTGAGAGGACACTCCCGTTTTCGGTGCCATCAGTCCCC
241-G---A---A---T---A---T---TGA-----
282 CGTC...TACAGCTCCCCAGCTCCCCCACCCTCCCCCACTCCCAACCAC
286 T---CATA-T---T---TGA---TTTA---A---A---T---
329 GTTGGGACAGGAGGTGTGAGGCAGGAGAGACAGTTGGAATCTTTAGAGA
333 TGA---T---T---T---A---TA-G---C---GG---
379 AG...ATGGATATGACCAAGTGTGCTATGGCTGTGCGATCCCAAGGCTGGT
383 --ACT---GT-G---AT-A---CACTA---T---A-T---
426 GGCTCAAGTCTGGCCCCACACAGCCCCAATCCAAACTGGCAAGGACGC
432 A---A-GA---T---A---A---A---TAT
476 TTCACAGGACAGGAAAGTGGCAGCTGTCTGCTCCAGCTCTGGCATGGTA
481 --TG--AA-A-A-----A-----A-----
526 GGAGGGGGAGTCCCTTGAACACTGCG...GTGTAGACTGGCCTGAACCACA
530 ---A-T---A---A---GA---CT-A---A---A---G---
575 GGAGAGGATGGCCAGGGTGGAGCTGGCATGCTCAATCTCAAGGAGC.T
576 -A---T-----A---A---T---CA---TG---A---C-T-C-
624 CCTCCAAAGGGTGGCGCTAGAG...GCCATGGAGGCAAGTGGACAAGGT
626 --C-T-GAA-A-AAT---A---AGGT-----C-----A---
670 GCAGGCAGGCTGGCTGGGGTCCAGGCCGAGACAGCAGGGGGTGAAGA
676 A---A---A---A---A---A---CT---TG-CA-A-AACA--
720 GGGATTCTTAATCACTCAGAGCAGTCTGTGACTTAGTGACAGGGGAGGG
724 -----ACTAG-A-----
770 GGCAAGGGGGAGGAGAGAAATGTTCTTCCAGTTACTTTCCAAATCTC
744 ...T---A---A-A-G---A---A---A---G---
820 CTTTAGGGACAGCTTAGAATTATTGCACTATTGAGTCTTCATGTTCCCA
791 -----
870 CTTCAAAACAAACAGATGCTCTGAGAGCAAACTGGCTTGAATTGGTGACA
841 -----A-----
920 TTTAGTCCCTCAAGCCACCAGATGTGACAGTGTGAGAACTACCTGSAAT
891 C---A---A---CA-G---C---C---T---C
970 TGTATATATACCTGCGCTTGTTTAAAGTGGGCTCAGCAGATAGGGTTCC
939 -----A-----
1020 CAGGAAGCTCCGAAACTCTAAGTGTGTGCTGCAATTTTATAAGGACTTCC
987 -----A-----
1070 TGATTGGTTTCTCTCTCCCTTCCATTTCTGCTTTGTTCATTTCATC
1037 -----C-----CTCGT-----T---CCT-C---G
1120 CTTTCACITCTTTCCCTTCTCTCGTCTCTCTCTCTAGTTTCATCCCTT
1087 -----T---C---AG-----T-----G-TT---
1170 CTCTTCCAGGCAGCGCGGTGCCCAACC.....ACACTTGTCT
1122 --G-----T--A---TG---ACATGGTTACCTA---GCA---
1207 GGCTCCAGTCCCCAGAACTCTGCTGCTCTTGTCTCTCTGCTGCCAGTA
1172 A---G---T---G-T---T---AA-T---CAT---
1257 CCAGCCCCACCCTGTTTTGAAGCCTGAGGAGCCTTGGGCTCTGCTGAGT
1221 ---GT-T-A-A---A---T---C-A-AGC---TT---C
1307 CCAACCTGGCCTGTCTG.TGAAGAGCAAGAGAGCAGCAAGSTCTTGCTCT
1267 --TC---AA---C-G---AA---AG-T-G---C---
1356 CCTAGGTAGCCCCCTCTTCCCTGGTAAGAAAAA..GCAAAAGGCATTTC
1317 -----C-----TGT-----A
1404 CACCCTGAACAACAGACGCTTTTACCCCTTCTACTCTAGAGAAGTGGACTG
1345 G---T-A---T-----T-----A-----
1454 GAGGAGCTGGGCCCCGATTGTGAGTTGAGGAAAGCACAGAGGCTCTCTGT
1394 --AA-T---T-AC---G---C---A-G-GAT-----A---
1504 GGCCTGCC...AGTCATCGAGTGGCCCAACAGGGGCTCCATGCCAGCGAC
1443 -----AG---G-T-CA-----CA-TAC---
1552 CTTGACCTCACTCAGAAGTCCAGAGTCTAGCGTAGTGCAGCAGGGCAGTA
1493 -----CT-----A---ATA-T-----T-----
1602 GCGGTACCAATGCAGAACTCCCAAGACCCGAGCTGGGACCACTACCTGGG
1537 ---G---TG-CAG---CAGATGC---TA---A---GTGAC---A---
1652 TCCCGCAGCCCTTCTCTGCTCCCCCTTTTCCCTCGGAGTTCTTCTTGAAT
1585 C---TTG-----ACAAA-A-ACA-ATC-CA---CT-T-CT-G-
1702 GGCAATGTTTGTCTTTTGTCTGATGCAGACAGG...GGGCCAGAACCA
1635 --A-T-C---AAC---C---AC---GG---T---GAA-CAT---C---T-C
1749 CACATTTCACTGTCTGTCTGTTCCATAGCTGTGGTATAGGSGCTTAGAGG
1685 ---T-----AAC---C---AG-ACA---GT-AAAT---T---A
1799 CATGGGCTTGTCTGGGTTTAAATTGATCAGTTTTCATGTGGGATCCCA
1731 -T---ACT-----G-----G-----A-T-----

1849 TCTTTTAACTCTGTTCAGGAAGTCTTATCTAGCTGCATATCTTCATC
1781 AAG-----CA-----
1899 ATATTGGTATATCTCTTTCTGTGTTTACAGAGATGTCTCTTA..TATCTA
1831 -----A---TC---G---
1947 AATCTGTCCAAGTGAAGTACCTTATCAAAGTAGCAATGAGACAGCAG
1881 -----
1997 TCTTATGCTTCCAGAAACCCACAGGCATGTCCCATGTGAGCTGTGCCC
1927 -----C-----
2047 ATGAAGTGTCAAGTGTGTGTGTCTGTGTATTTCAGTTATTG.TCCCTG
1977 -----G---A---TC---AC-T---CA
2096 GCTTCTTACTATGTGTATATCAAGGAGTGAACATCATAGAAACTG
2027 -----C-GC---G-A---TC---
2146 TCTAGCACTTCTTGGCAGTCTTTAGTGATCAGGAACCATAGTTGACAGT
2077 -----G---C-----G-----
2196 TCCAATCAGTAGCTTAAGAAAAAACCGTGTGTCTCTCTGGAATGGTT
2127 -----TGA---A---A---A-----A-----
2246 AG...AAGTGAGGGAGTTTGGCCCGTCTGTGTTGTAGAGTCTCATAGTT
2177 --AACT---A---TC---CC---A-----
2292 GGACTTTCTAGCATATATGTGTCCATTCTCTATGTGTAAAAGCAAGTC
2225 -----TG---T---A-----G---A-----AC-
2342 CTGCAACCAAACTCCCATCAGCCCAATCCCTGATCCCTGATCCCTTCCAC
2274 -----GCT-T-TG---G-----GT---TTG-
2392 CTGCTCTGTGATGACCCCCCAGCTTCACTTCTGACTCTTCCCCAGGAA
2308 -----TT---T---T-TT---TC-TG-----GT---C---TG-AG-
2442 GGGAGGGGGTCCAGAGAG.....AGGGTGAGTCTCTCC
2358 -----G-A-----GAATCTGGAGGATCC---A---AT---T-
2476 AGAAGT...CTTCTCCAAGGACAGAGGCTCTGCCCCCATAGTGGCC
2408 T-----CCTG---T---GT---A-----C---A-TG---CC---
2522 TCGAAGT...CTGGCACTACCAAGGACACTTATCCA.CGAGAGCGCAG
2454 C-A-C-GCC--AT---C-TCT---C-T---TACTTTT-A---A
2568 CATCCGACCAGGTTGTCACTGAGAAGATGTTATTTTGGTCAG.TTGGGT
2504 --C-TAG-----A---C---T-----G---T-C---AA
2617 TTTTATGTATTA...TACTTAGTCAAAATGTAATGTGGCTTCTGGAATCA
2551 -----CC-G---GGCT-----C---CA-----
2663 TTGTCCAGAGCTGCTTCCCCGTCACTGGGCGTCACTGGTCTGTTAAG
2586 -----AC---AA---TGGG-ATCCC-----T-G---
2713 AGGAGTGGCTGGCCCCACAGGCCCTTGTCAACCATGACAGTTCATTCA
2619 -----AC-T-----
2763 GGGCCGATGGGGCAGTGTGTGTTGGGAACACAGCATTTCAAGCGTC.ACT
2626 A---T---TA-AT-T-TACTT-----TTG---C---TGG-T-C-GTT-
2812 TTATTTCACTTGGGGCCCCACCTGCAGCTCCCTCAAAGAGGCGATGTCCCA
2676 ---T-C-T---TTTT-T-----CCAC--CCAC-A-C
2862 GCCTCTTTCCCT...TCCAGTTTATTCAGAGCTGCCAGTGGGG...C
2724 C---CAC---A---GTATG-AG-AC-G---T---A---A---AG---T-GTAA
2904 CTGAGGCTCCTTAGGGTTTCTCTCTATTTCCCCCTTCTTCTCTATTCC
2774 ---CA---T---C---TTG---CT---
2954 CTGCTCTTTCCCAAA...GGCATCAGAGTCACTGCGCTTTTCCAGCAGGC
2822 ---A---A---TAAG-----A-----T-----A-----
3000 AGCCTTGG.CGGTTTATCGCCCTGGCAGGCGGGCCCTGCAGCTCTCAT
2869 ---T---TG-G---T-TC-----A---C-----A-----G
3049 GCTGCCCCCTGCCTTGGGGTCAAGTTGACAGGAGTTGGAGGG.AAAGCCT
2913 -----T-----T-----A-----
3098 TAAGCTGCAGGATCTCACCAGCTGTGTCCGGCCAGTTTGGGGTCTGA
2963 ---TCATG---T---T---ACC---AA-G---
3148 CCTCAATTTCAATTTTGTCTGTACTTGAACATTATGAA..GATGGGGGCC
3013 -----TT-----T.....G.A..GT.TG..T..TA
3196 TCTTTCAGTGAATTTGTGAACA..GCAG.AATTGACCGACAGCTTTCCAG
3056 C---AA-C-----A---G-ATC---A---C-----A-----AGA
3243 TACCATGGGGCTAGGTCAATTAAGGCCACATCCACAGTCTCCCCACCCCT
3106 -----C-----T-----T-----
3293 TGTTCAGTTGTTAGTTACTACCTCTCTCTCTGACAATACTGTATGTCTG
3151 -----C-----TC-CAGAT-G-T-G-----A---C---
3343 CGAGCTCCCCCAGGTCTACCCCTCCCGGCTGCTGCTGGTGGGCTTG
3201 --C-A---...AT-A---AT---TGG---AAG-T-----T-CT-A-TC---
3393 TCATAGCCAGTGGGATTTGCGGCTTGTGACAGCTCAGTGAAGCTGGAGATAC
3246 --T-A-----A-C---T-CC---GT-TG-----T---AG-CT-
3443 TTGGTCACAGCCAGGCGC...TAGCAGAGTCCCTTCTGTGTGATGTCTGA
3295 -----T-TAG---T-----T-----C-G-----
3490 TTCCATATCAAAAGGCACAGGGGACCCAGAAAGCCCACTCCCCCAA
3345 ---T---T-----G-T-T---T-----TC-----G
3540 TCCATCAGTGCCAAACTAGCCAAAGCGCCCGAGCTTCTCAGCTCGCTGGAT
3395 -----A---C-T-AT-----A-----
3590 GCGGGAAGCTGCTACTCGTGAGCGCCAGTGGGGTGACAGCAATCTTCTG
3430 -----A-----A-----C-----
3640 TTGGGTGGCATCATTTCCAGGCCGGAAG.CATGAACAGTGCACCTGGGACA
3447 --CA---CT---CAA---A-TA-G-C-----T-G-G---AC

Fig. 3 (1)

12/21

3689 GGGAGCAGCCCCAAATGTACCTGCTTCTCTGCCAGCTTTTCATTGCT
3497 CA-----A---G---T-----G---G---T---ATT-----C-G-----
3739 GTGACAGTGATGGCGAAGAGGGTAATAACAGACACAACTGCCAAGTT
3542 C-A-----A-----C-A---G-----TTA-A---G-----G-----
3789 GGGTGGAGAAAGGAGTTTCTTTAGCTGACAGAACTCTGTAATTTTAAATC
3589 T--CA-----AA---C-----C-C-----G-----A
3839 ACT..TAGTAAGCGGCTCAAGC...CCAGGAGGGAGCAGAGGGATACGA
3639 GG-TG--A---T---T---C-A---CCAT-----AAA--A-A-G-TA--
3883 GCGGAGTCCCTCTGGCGGGACCATCTGGAATTTGGTTTACGCCAAGTGGAG
3689 A-T-----CAGAT-A-----G-CT--T--CA-GC-----A-----
3933 CCTGACAGCCAGAACTCTGTGTCCCCGCTCTAACACAGCTCTCTTTTCCA
3734 T-CAG-T-T-T-----CAG-A-TTC-----T-----C-C-T--
3983 GAGCATTTCCAGTACGGCTCTCTGGGCTGACTGGGCGAGGGAGGTACAG
3779 -----C-----TCA-----A-----A--CA-C-----
4033 GTACCACTTTTAAAGATCTTTTGGGCATATACATTTTAGCTGTGT
3821 --G-----C-----CT-----T-----CAG--CA-----A-A-
4083 CATGGCCCCAAATGGATTCTGTTTCAAGTTTCAACCTTCAGATTTCTAGG
3869 -T-----A-----GTA-A-----G---A-----A-----
4133 ACCTGTGTCTTAGACT.....TCAGGGAGTCACTGTTTCTAG
3914 G-AGA--A-TG-----CAGAAAAAAGCC-CT-TG-A-T-TG--A-A-AGC-
4171 AGTTCCTACCATGGAGTGGGTCTGGAGGA.....CTGCCCCGTTGGGG
3964 -AG-A-A-AA-CT--AG---G--A-G-C-GATGCCG-A---TCA--CCA
4214 GGGCAGAGCC..CTGCTCCCTCC.....GGGTCTTCTTACTCT
4014 -A-CA-----CT-----A--ACATCCTTTTCT--C-C-T--T--
4250 TCTCTCTG.....CTCTGACGGGATTTGTTGATTCT
4063 -CTTCAGTGAACAGCCCCA-----A-----A-----
4281 CTCCATTTTGGTGTCTTTCTCTTTTAGATATTGTATCAATCTTTAGAAAA
4113 -----TA-----G--C-A-----
4331 GGCATAGTCTACTTGTATATAAATCGTTAGGATATGCTCCCCCAGGGTC
4155 -T-----A-----T-C-A-----
4381 TAAAAATACATATTAGAGGGGAAAGCTGAACACTGAAGTCAGTTCTCAA
4205 -----A-T-GC--A-----CT-----AC-----G-----
4431 CAATTTAGAAGGAAACCTAGAAAAACATTTGGCAGAAAAATTACATTTGA
4255 -----T-----AA--A-----T-----TAA
4481 TGTTTTGAATGAATACAAGCAAGCTTTTACAACAGTCTGTATCTAAAAA
4305 -----AG--G-G-----GA--A-A-T-----T-----
4531 TACTTAGCACTTGGCCTGAGATGCTGCTGAGCAATTACAGGCAAGGGGA
4351 -----TT-A-----G--T--AA-----TGAACCTGA-----T-----
4581TCTGGAGGTAGCCGACC
4450 GAGATTGAGGTGTTTTAGCATTTGGAAGGCCAC---TTG-----T-G-----
4598 TGAGGACATGGCTTCTGAACCTGTCTTTTGG.....GAGTGGTATG
4500 CC--A-CTA-----C-T-----A---AATGGAGGT-----C-----
4639 GAAGGTG.....GAGCG
4549 CC--A--CCAAAGCTGATGAGACCAAGCTCTTGGTTTATCAATTT--A-A-
4651 TTCACCACTGACCTGGAAGGCCAGCACCACCTCTTCCCACTCTTCTC
4599 C-----A-----A-AGTGT--G-----
4701 ATCTTGACAGAGCCTGCCCCAGCGCTGACGTGTCAGGAAAAACACCCAGGG
4639 T-----A-----A-----TGCACA
4751 AACTAGGAAGGCACCTTCTGCTGAGGGGAGCCTGCTT..GCCACTCC
4673 --T-----G--GT--A--T--ACA-A--T--CT--C-GG--C-----
4799TGCTCTGCTCGCTT.....CGGA
4723 CATCTTTTG-A-----A-CTA---GACCTTACGATCTTTGGCACATA--A-
4817 TCAGCTGAG.....CCTTCTGAGCT.....GG
4773 ATG-----T-TAGCAAGCACTTTGGCATGC---C-A-A--TACCCAGAG-
4839 CCTCTCACTGCTCCCCAAGGCCCTTCTGCTGCTT.....
4823 -----C-----TT--C-AGT-----T-T-CTGGGAGGTGTTA
4875GTCAGGAGGAGGAGGAAGGAGCAGGTG
4873 GAGCCCATAGAAATGAGAGGAGAAA-AA-A-----A--G-C-G-----A
4900 TGAGGGCAGTGCAAGGAGGGAGCACAACCCAGCTCCCGCTCCGGGGCTC
4923 GT-AAAAG-CT-TG-----A-AG--G--T--TAGG-----
4950 CGACTTGTGACAGGAGGAGCCAGCCCTGGAGG.....AAATCCTACC
4960 -G-G-T-----A--GA-T--C-----T-T-GAAT--G-G-C-T--
4995 TTTGAATTCAGAACATTTGGGGAATTTGGAATCTCTTTTCCCCCAAC
5005 --G-G-T-----AC-G--C-----GA-C--TG-T--TTG-
5045 CCCCATTTCTGCTACCTTTAATCAGGTCTGCTCAGCAGTGAGAGCAGA
5055 T-----TC-----T--GG--C-----T--C-A--A-----
5095 TGAGGTGAAAAGGCCAAGAGGTTTGGCTCTGCCCCACTGATAGCCCCCT
5098 --C-C-----A-----T-----T-----
5145 CCCCGCAGTGTTTGTGTCTCAAGTGGCAAAGCTGTCTTCTGCTGAGCC
5147 --T-----T-----
5195 TGATTATATCCAGTAACACATAGA...CTGTGCGCATAGGCTGCTTTGT
5197 -----G-----TT--GATT--AT-----T-----A
5242 CTCTCTATCTCTGGGCTTTTGTCTTTTGTCTTTTGTCTTTTGTCTTT
5246 --T-----T-----A-----C-----
5292 TCTGTCCCTTTTATTTAAGCACCAGCTAGACACACAAAGCAGTTGAATT
5296 C--A-----T-T-----A-----

5342TTTATATATATATCTGTATATGCACAATATAAACTC
5343 TATATATATATATA-----T-----
5380 ATTTTGTCTGTGGCTCCACACACACAAAAAAG...ACCTGTTAAAAAT
5393 -----A-G-----C-----AAAA--T-----
5426 ATACCTGTGTCTTAATTACAATATTCTGTATAACCATAGCATAGGACAAG
5443 -----AG-----
5476 GGAAAAATA.AAAAAAGAAAAAAGAAAAAAGCAGAAATCTGTCTGC
5493 -----A-TTT-----A--AA--C-----
5525 TGGTCACTTCTTCTGTCCAAGCAGATTCTGTGCTTTTCTCTGCTTCTTT
5543 -----AA-----CT--A-----G-----
5575 CAAGGGCTTCTCTGTGTCAGGTGAAGGAGGCTCCAGGAGCAGCCAGGTT
5592 --A-A--C-----T-A-----A-----
5625 TTGCACTCTGTCTTCTCCGCTGTGTGAAAGAGGTCCCAAGGTTCTGGG
5642 --TG-------TC-------G--C-----A-----
5675 TGCAG.....GAGCGCTCCCTT
5690 -A--GACAGTTCAATTTCAGCATGGGCTCAGGAGACAA--A-----
5692 GACCTCTGAAGTCCGGAACGTAGTCCGACAGCTGGCTGCCCTTCCACC
5740 T--A-----C--A-A--T-----G--A-T-----A-----C-----
5742 TCT.....GGGAGCTGGAGTCCAGTGGGCTGACTCCCCAGTCT
5786 --AGGATG-----A-CA-G--TG-----A-CA-----T-----
5785 CCCTTCCCGTGACCTGGTCAAGGTGAGCCATGTGGAGTCAGCCTTCGAG
5833 A---T-----C-----AA-----TG--AC--A--T--GA--A
5835 GCCT..CCCTGCCAGTAGGG..TCCGAGTGTGTTTCACTCTCC..CACTCT
5878 --T-TC--A-TT--C-----C--A--CA-T-AAA-A--C--AA--T--
5881 CTGAGCTTGGGGCTGGAGCGGAGAGCGGAGGCTGGCTGTCTCGGA.
5928 -GT--T--A--A--A--A--A--A--TA--A--ATT--T--A--CA-G
5930 ACCTGTGAGCTGCACCAGGTAGAAGCCAGGAGCCAGAAATCATGTGCG
5978 GTG--A-----T-----T-----C--CA-C-A
5980 TCACTCCAAGGGTCCCTCCAG..GAGTAGTGAAGACTCCAGAAATGTCC
6028 A-C-T-----T-----AA-----CA-----
6029 CTT..TCTTCTCCCTCATCTACAGTAATGTCAATTGCTTTTGTAAATTC
6078 --TC-----TGC--C-T-----C-----
6077 TTAATGAGCAATATCTGCT...AGAGAGTTTAGCTGTAACTTCTTT
6128 -----T-----AAAA--A-AA-----
6122 TTG...ATCATCTTTTAAATAATTAGAAACACC.....AAAA
6178 -----CAAA--GG--A--C--A-----A-----CCCCCCCCAA--
6158 AAATCCAGAACTTGTCTTCCAAAGCAGAGAGCATTATAATCACCAGGG
6228 --C-----
6208 CCAAAAGCT..TCCCTCCCTGCT.....GTCAATTGCTTCTTCT
6275 -----T--G-----A-A-CT--ACCCCATCTCTCTCA-G-----G-----
6244 GAGGCTGAATCCAAAAGAAAAAGCAAGGCTTTTCACTGGCGGGG
6325 A-----A-----C-GC--G-T-TTTG-----GG--A-T-----
6294 CTACCCGTGAGCCCTTCCGAGGAGCAGGCTGGGGCAGCTCTGGGCCCA
6373 T-----GAG-TC-T--A--T-----A-----A-A--C
6344 CATCC..GGGCCCAGCTCCGGCTGTGTTTCACTGTAGCAGTGGGTCAATG
6421 T-----TA-----C-TT-----TA-G-A--AA-----A--T--CA
6392 ATGCTCTTTCCCAACCCAGCCTGGGATAGGGGAGAGGAGGAGGAGGCG
6471 -CAT-A--C-----G--G--A--A--G-----A--A--C-AAG--TT
6442 GTTCCGCTG...ATGTTGGCCGTGAACAGGTGGGTGTCTGCTGCTGCT
6521 --CT--A--CTACT-AC-----ACTG-----A--A-CAT--AT--
6488 CCAGTGTGCTTTTCTGACTGACATGAATTCAGGCGGAGTTAGCCTC
6571 --TA--A-----A--G-----TG--G--AG--A--C-----T-----
6538 ACCGGTGACCTCTAGCCTGCGGATGGAGGGGGGCCACCCGGTTCA
6621 --T-----A-----T-----A-----T-----
6588 GTGTTTCTGGGAGCTGGACAGTGGAGTGCAAAGGCTTGCAGAACTTGA
6669 A-----C-----A-----A-----
6638 AGCCTGTCTCTTCCCTTGTCTACCAAGGCTCC..TTCCGTTTGAATTGT
6719 --TT--A--T-----G-----T-----A-----
6687 ACTGCTCAATCAATAACAGCCGCTCCAGAGTCACTAGTCAATGAATATA
6769 -----G-----TG-----
6737 TGACCAAAATATCACCAGGACTGTACTCAATGTGTGCGGAGCCCTTCC
6819 -----C-----T-----T-----
6786 CATGCTGGGCTCCC..GTGTATCTGACACTGTAACTGTGCTGTGTTTGC
6869 TG-----T-----C-----T-----
6835 TCCCTTCCCTTCTCTTCTTGTGCTTTTCTGGGTTTCT
6919 --T-----T-----C-----A-----
6885 TGTTTGGGTTTGGTTTGGTTTATTTCTCTTTTGTGTTTCAAAATGA
6969 -----T-----
6935 GGTCTCTCTACTGGTCTCT..TTAAGTGTGTTGAGGCTTATATTGT
7017 -----T-----C-----
6984 GTAATTTTGGTGGGTGAAAGGAATTTGCTAAGTAAATCTCTCTGTGT
7067 -----C-----
7034 TTGAAGTGAAGTCTGTATTGTAAGTATGTTTAAAGTAATTGTTCCAGAGA
7117 -----A-----
7084 CAAATATTCTAGACACTTTTCTTTACAAACAAAAGCATTCCGAGGGAG
7167 -----GC-----GT--A-----A-----T-A-----

Fig. 3 (2)

```

7134 GGGGATGGTGACTGAGATGAGAGGGGAGAGCTGAACAGATGACCCCTGCC
7216 --AAG-----A-A---CA-----CA-T-----T---T-----
7184 CAGATCAGCCAGAAGCCACCCAAAGCAGTGGAGCCAGGAGTCCCACTCC
7263 -----T---A-----TA---A-T--TT-T---T---
7234 AAGCCAGCAAGCCGAATAGCTGATGTGTTGCCACTTTCCAAGTCACTGCA
7310 --T--AG-GA-T--T-----T-----T---A---AA---
7284 AAACAGGTTTGTGTTCCGCCAGTGGATTCTGTGTTTGTCTCCCTCCCC
7358 -----C-----A---A-----T---T-----
7334 CCGAGATTATTACCACCATCCCGTGCTTTAAGGAAAGGCAAGATTGATG
7401 -----T-----G-A-----
7384 TTTCTTGAGGGGAGCCAGGAGGGGATGTGTGTGTGCAGAGCTGAAGAGC
7422 --AA-CT--A-T-----A-A--A--A-TA-C-...C-----
7434 TGGG....GAGAATGG...GGCTGGGCCCCCAAGCAGGAGGCTGGG
7465 --T-CTCACT--T---AAA---T--T-TGAGTTT-----A--AC
7475 ACGTCTCT.GCTGTGGGCACAGGTCAG..GCTAATGT.....TGGC
7515 C-A--G-G-ACA---G-G-A-A---AA-A---AT-AGCCGCTCCC--C-
7512 AGATGCAGCTCTTCTCGGA.CAGGCCAGGTGGTGGGCATT.CTCTCTCCA
7565 TA-GAT-C-----AA-A--TA--T-A---CCA--A---AT-G---A--
7560 AGGTGTGCCCCGTGGGCATTACTGTGTTAAGACACTTCCGTCACATCCAC
7615 ---CA--TTT-AAA-A---G--CAG-C-G-----T---T-CT---T-
7610 CCCATCTCCAGGGCTCAACAC...TGTGACATCTCTATTCCTCCACCTC
7665 GTTGC--C-T--TA-A--GT--TAA-C---T-----G-----
7657 CCCTTCCAGGGCAATAAAATGACCATGGAGGGGGCTTGCACTCTCTTGG
7708 G--A-G--T---GG-----TAGCA--ACTC---T-----CA
7707 CTGTCAACCCGATCGCCAGCAAACTTAGATGTGAGAAAACCCCTTCCCAT
7753 A--G-T-TA--TC--A-----TC-G-GCC-----T-C-A-----GT-
7757 TCCATGGCGAAAACATCTCCTTAGAAAAGCCATTACCCTCATTAGGCATG
7800 --T-----A-C-...---C--T---TG---TT-----GCAG--T
7807 GTTTTGGGCT.....CCCAAAACACCTGACAGCCCCCTCCCTCTCTG
7845 ---CCA-C--AATGTAAGAGG--C-G-G-A-TGTT--T-GGAG-----
7849 AGAGGCCGAGAGTGCTGACTGTAGTGACCA.TTGCAATGCCGGGTGCAGCA
7893 ..T-T--C--T--AC-----G---GC-ATA-TAGTT-TT-
7898 TCTGGAAGAGCTAGGCAGGGTGTCTGCCCCCTCCTGAGTTGAAGTCATGC
7941 G-----C-A-C--A-----ACA-----AA-A-CC-T--TG-A-
7948 TCCCTGTGCCAGCCAGAGGCCGAGAGCTATGGACAGCATT...GCCAG
7991 -----C-----G--GG-T--A-C--T-G-AT-G-GCA-----
7995 TAACACAGGCCACCTGTGTCAGAAGGGAGCTGGCTCCAGCCTGGAAACCT
8040 ...G--A-----T-----TT--T-A-----TCAA---TC
8045 GTCTGAGGTTGGGAGAGSTGCACTTGGGGCACAGGGAGAG.GCCGGGACA
8079 A--A-----CA-...GACAA--G--A--A--AT--A-----
8094 CACTTA....GCTGGAGATGTCTCTAAAAGCCCTGTATCGTATTACCT
8128 G--C--GCTGG-----GG---TG-----C-G-----
8139 TCAGTTTGTGTTTGGGACAACTTCTTAGAAAATAAGTAGGTCGTTT
8178 -----C-----
8189 TAAAAACAAAAATTATTGATTGCTTTTTTGTAGTGTTCAGAA.AAAAGGT
8228 -----A--C-----
8238 TCTTTGTGTATAGCCAAATGACTGAAAGCACTGATATATTAAAAACAAA
8276 -----
8288 AGGCAATTTATTAAAGGAAATTGTACCATTTCAGTAAACCTGTCTGAATG
8326 -----
8338 TACCTGTATACGTTTCAAAAACACCCCCCCCCCACTGAATCCCTGTAACC
8376 -----A-----C-----
8388 TATTTATTATATAAAGAGTTTGCCTTATAAATTTA
8422 -----

```

Fig. 3 (3)

dashed line: putative promoter

full line: sequence-conserved high-energy sequence

14/21

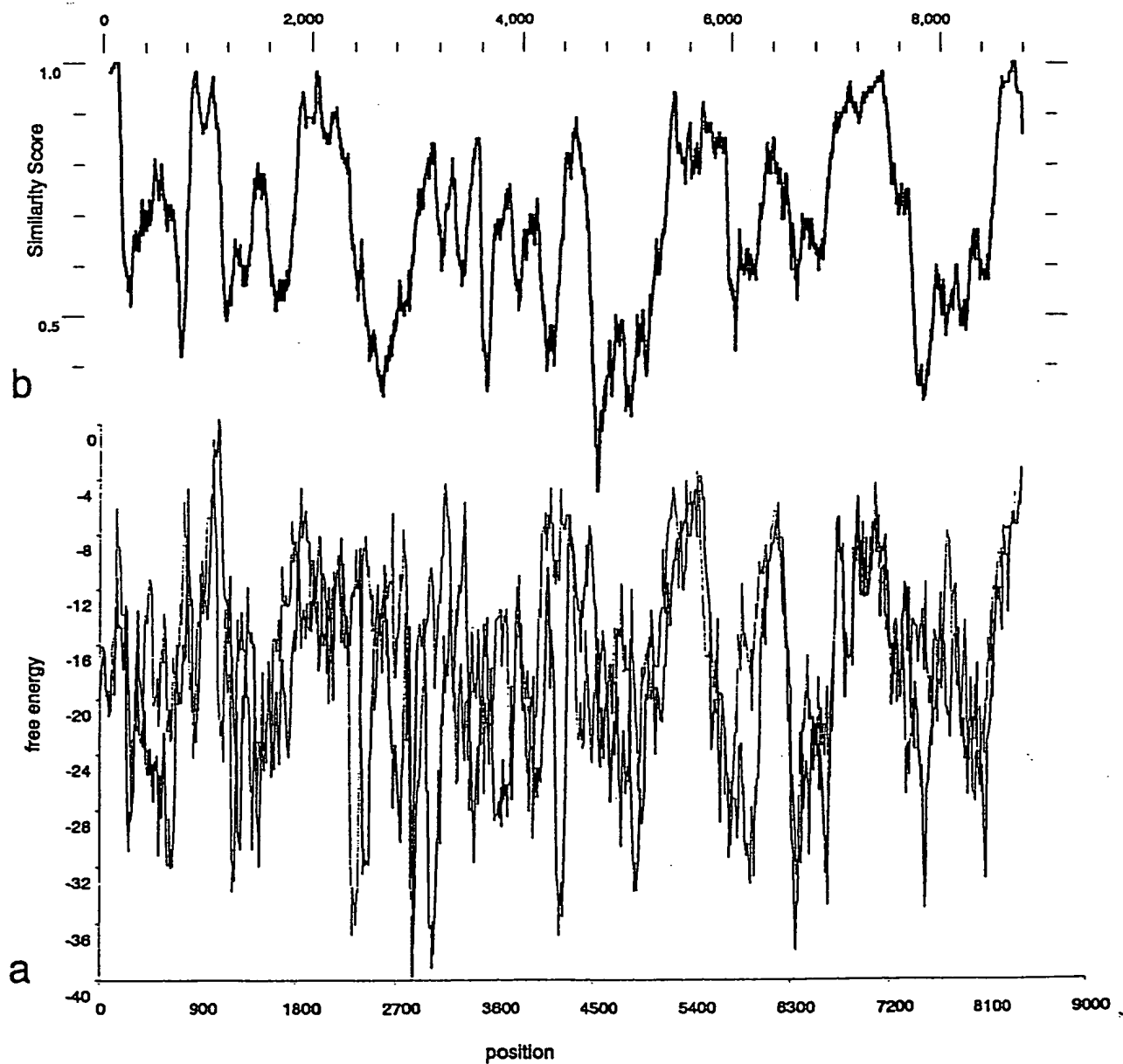


Fig. 4

1
human TTGCTGCAGATACTACTGACCAGACAAGCTGTTGACCAGGCACCTCCCTCCCGCCAAACCTTT.....CCCCATGTGGTCGT
schim
orang
makak
hamst
mouse
rat
kaeng

101
human TAGAGACAGAGCGACAGAGCAGTTGAGAGGACACTCCCGTTTTCGGTGCCATCAGTGCCCGTCTACA...GCTCCCCCAGCTCCCCC...ACCTCCCCC
schim
orang
makak
hamst
mouse
rat
kanga

201
human ACTCCCAACCACGTT.GGGACAGGGAGGTGTGAGGCAGGAGAGACAGTT..GGATTCTTTAGAGAAGA...TGGATATGACCAGTGGCTATGGCCTGTGC
schim
orang
makak
hamst
mouse
rat
kanga

301
human GATCCCAACCGTGGTGGCTCAAGTCTGGCCCCACACCAGCCCCAATCCAAAACCTGGCAAGGACGCTTCAAGGACAGGAAAGTGGCACCTGTCTGCTCC
schim
orang
makak
hamst
mouse
rat
kanga

401
human AGCTCTGGCATGGCTAGGAGGGGGAGTCCCTTGAAGTACTGGGT.GTAGACTGGCTGAACACAGGAGAGGATGGCCAGGGTGGGTGGCATGGTCC
schim
orang
makak
hamst
mouse
rat
kanga

501
human ATTCTCAAGGGACG.TCCTCCAACGGGTGGCGCTAGA....GGCCATGGAGGCGTAGGACAAGGTGCAGGCAGGCTGGCTGGGGTCAGGCCGGGGCAG
schim
orang
makak
hamst
mouse
rat
kanga

601
human AGCAGCGGGGTGAGAGGATTCTAATCACTCAGAGCAGTCTGTGACT.....TAGTGACAGGGGAGGGGCAAAGGGGGAGGAGAG
schim
orang
makak
hamst
mouse
rat
kanga

701
human AAAATGTTCTTCCAGTTACTTTCCAATTCT...CCTTTAGGGACAGCTTAGAATTATTGCACTATTGAGTCTTCAT...GTTCCCACTTCAAAACAAA
schim
orang
makak
hamst
mouse
rat
kanga

801
human CAGATGC....TCTGAGAGCAAACTGGCTTGAATTGGTGACATTTAGTCCCTCAAGCCACCAGATG....TGACAGTGTGAGAACTACCTGGATT
schim
orang
makak
hamst
mouse
rat
kanga

901
human GTATATATACCTG
schim
orang
makak
hamst
mouse
rat
kanga

Fig. 5

Partial sequence of the non-coding RNA gene from hamster

1 TTGCTGCAGA TACTACTGAC CAGACAAGCT GTTGACCAGG CACCCCCCA
51 ATACTCCCCC AATGTGCTCA TTAGAGATAG CAGTTGAGAG GACACTCCCA
101 TTTTGGGTGC CCTGTCCATA GCTTCCCTGA CTCTTCCACC ACCCCAAGTC
151 CCAATCTGAG GGACCGGGAG GTGCGAGGCA GGAAAAATAT TGGATTCTTT
201 AGAGAAGACT AGAGGTGACC AGTGACTGTG GCCCAGTAAT TAGAACTGTG
251 GTGGCACAAG TCTGGCCCCA CATCCACCCA ATCCAAAAGT GATAAGGATA
301 TTTTGAAGAA CAGGAAAGCA GTACCTGTCT GATCCAGCTC TGGTATAGGT
351 AGGAGTGAGT CCTGAACTGC TGGATTACAG ACTGGCTTGA GCCACAGAAG
401 ATGATGGACC AGAGTAAAGT ATCATCACCT GTCACAAGG CATGCTTCAC
451 TAGAGAATAA TTCTAAAGAG GTGCCATGGA GGCAGCAGGA CAAGGCACAA
501 GCAGTCTGGG TGGGGGTCAA GCCAGACCTA GTGCCACAGA ACAAGAGAGC
551 AATCTGTGAC TAGTAGTTAG GGACTTTGTG GATGGGACAA GGGGCATGGG
601 GGAAGAAATG AAAATATTCT TCCAATTACT TTCCAGTTCT CCTTTAGGGA
651 CAGCTTAGAA TTATTTGCAC TATTGAGTCT TCATGTTCCC ACTTAAAAAC
701 AAACAGATGC TCTGAAAGCA AACTGGCTTG AAATGGTGAC ACTTTGTCCC
751 ACAAGCCACC AAATGTGGCA GTGTTTAGAA CTACCTGGAT CTGTATATAC
801 CTG

Fig. 5a

17/21

Partial sequence of the non-coding RNA gene from kangaroo

1 TTGCTGCATA TACTACTGAC CAGACAAGCT GTTTATCAGG CTTTTTAGGG
51 TACACCAGCA CCTGCCCTCC ATTCATCCCT GTTGGGAGAG GGATGGTGTA
101 CTGGTTGTCA CTAGAGACCT AACAGAGTAG GGTTAGTGGG AGCTTACATT
151 TTCAGTGCCA TTAACATTCT AGTCCAAGGT CTTAAATTAT TATGTTGAGG
201 GGTTTTTTTTT CCCCTGAGGG GGCCGGGGGG TGGGGGGAGG GTTGATTAGA
251 TTCCTTAGGA AAGAGGGTTG AGACAGACAG CAGAGCACTG AGCAGTTGGC
301 ACTAAAGGAG ACCTTGACTA GGGGCCAGGT GGCATCATCT AATCCCAAGG
351 GGCTCCAAGT GAGTATTAGG GTGGGGGAAG ACATTATAGA AGGAATAGAA
401 ACAGGATAGC TCAGCCTAAA GAAGAGCGGT TAAAACCCTA CCCACCAGGA
451 GTTGACTTGA AAGAGGCCCC TATGGAGGAA TCCCCAACCA CCAAAGCAA
501 TCTTGAGCTG CAGCTGCTTC ATTTAGTGGA CCTTGTGTAT ATCTGGGTGT
551 GTATGCACAT AGATAGACAG TGAGAAAGAA AACTGTTCTT CCAGTTCTTT
601 TCCAGTGCTA CTAGCTTAGG GACAGGTTAG AACTGTCTGC ACAATTGTGT
651 GATCATTCCC ATTCCCACTT CAAAACAAAC TGA CTGAGAT GTTCAACAGA
701 AAAGTGGCTT CAATGGGTAA CATGCCCTTG CCACCTACTT AAGACACTGG
751 TGTGATGGGG TTTTGAATC CCTATATTG TAGGTATCTG

Fig. 5b

Partial sequence of the non-coding RNA gene from makaka

1 TTGCTGCAGA TACTACTGAC CAGACAAGCT GTTGACCAGG CACCTCCCCT
51 CCCGCCCAAA CCTTTCCCCC ATGTGGTCGT TAGAGACAGA GCAGTTGAGA
101 GGACACTCCC GTTTTCGGTG CCATCAGTGC CCCGTCTACC ACTCCCCCAG
151 CTCCCCCACC CTCCCCCACT CCCAACCACG TTGGGACAGG GAGGTGTGAG
201 GCAGGAGAGA CAGTTGGATT CTTTAGAGAT GGATGTGACC AGTGGCTATG
251 GCCCGTGCGA TCCCACCCGT GGCGGCTCAA ATCTGGCCCC ACCCCAGCCC
301 CAATCCAAAA CTGGCAAGGA CGCTTCACAG GACAGGAAAG TGGCACCTGT
351 CTGTTCCGGC ATGGCTAGGA GGGAGTTGTC CCTTGAAC TA CTGGGTGTAG
401 ACTGGCCTAA ATCACAGGAG AGGATGGCCC AGGGTGAGGT GGCATGGTCC
451 ATTCTCAAGG GACGTCCTCC AGTTGGTGGC ACTAGAGAGG CCATGGAGGC
501 AGTAGGACAA GGCACAGGCA GGCTGGCCCA GGGTCAGGCC GGGCCGAACA
551 CAGCGGGGTG AGAGGGATTC CTCGTCTCAG AGCAGTCTGT GACCGGTAGT
601 TAGGGACTTA GTGGACAGGG AAGGGGCAA GGGGGAGGAG AAGAAAATGT
651 TCTTCCAGTT ACTTTCCAAT TCTACTCCTT TAGGGACAGC TTAGAATTAT
701 TTGCACTATT GAGTCTTCAT GTTCCCCTT CAAAACAAAC AGATGCTCTG
751 AGAGCAAACCT GGCTTGAATT GGTGACGTTT AGTCCCTCAG GCCACCAGAT
801 GTGATGGTGT TGAGAACTAC CTGGATATGT ATATATACCT G

Fig. 5c

Partial sequence of the non-coding RNA gene from orangutan

1 TTGCTGCAGA TACTACTGAC CAGACAAGCT GTTGACCAGG CACCTCCCCCT
51 CCCGCCCAAA CCTTTCCCCC ATGTGGTCGT TAGAGACAGA GCAGTTGAGA
101 GGACACTCCC GTTTTCGGTG CCATCAGTGC CCCGTCTGCA GCTCCCCCAG
151 CTCCCCCAC CTCCCCACT CCCAACCACG TTGGGACAGG GAGGTGTGAG
201 GCAGGAGAGA CAGTTGGATT CTTTCGAGAA GATGGATATG ACCAGTGGCC
251 ATGGCCTGTG CGATCCCACC CGTGGCGGCT CAAGTCTGGC CCCACACCAG
301 CCCCAAATCCA AAACCTGGCAA GGACGCTTCA CAGGACAGGA AAGTGGCACC
351 TGTCTGCTCC AGCTCTGGCA TGGCTAGGAG GGAGTCGTCC CTTGAACTAC
401 TGGGTGTAGA CTGGCCTGAA CCACAGGAGA GGATGGCCCA GGGTGAGGTG
451 GCATGGTCCA TTCCTCAAGGG ACGTCCTCCA ACGGGTGGCG CTAGAAAGGC
501 CATGGAGGCA GTAGGACAAG GCGCAGGCAG GCTGGCCCGG GGTCAGGCCG
551 GGCAGGGCAC AGCGGGGTGA GAGGGATTCC TAATCACTCA GAGCAGTGTG
601 TGACTGGTAG TTAGGGACTC AGTGGACAGG GGAGGGGCGA GGGGGCAGGA
651 GAAGAAAATG TTCTTCCAGT TACTTTCCAA TTCTCCTTTA GGGACAGCTT
701 AGAATTATTT GCACTATTGA GTCTTCATGT TCCCATTCA AAACAAACGA
751 TGCTCTGAGA GCAAACCTGGC TTGAATTGGT GACATTTAGT CCCTCAAGCC
801 ACCAGATGTG AGTGTGAGAA ACTACCTGGA TTTGTATATA TACCTG

Fig. 5d

Partial sequence of the non-coding RNA gene from rat

1 TTGCTGCAGA TACTACTGAC CAGACAAGCT GTTGACCAGG CACTCCCCAC
51 AACAAACAACC CCCTCCCTCC TCACCCACACC CCTATCCCCT GTGTGCTCAT
101 TAGAGAGGGC AATTGAGAGG ACACTCCCAT TTTTGGTGCC ACTGATGCCC
151 TGTCCATAGC TTCCCTGACT TTTACACCAC CCCAACTCCC AATCTGAGGG
201 ACTGGGAGGT GTGACGCAGG AGAAACTATA TAGGACTCTT GGGAGAAGAC
251 TATAGAGTTG GCAAGTGATT GCGCCCCAGT AATTCCAAC TGGGTAGCAC
301 AAGTCTGGCT CCACACCAAC CCAATCCAAA ACTGACAAGG ACATTTTGCA
351 AAAAATGAAA GTGGCATT TG TCTGATCCAG CTCTGGCATG GCTAGAGATG
401 AGTCTTAAAC TGTGGCTTA TAACTGGCC TGAGCAACAG AAGAGGATGG
451 CCCAGAGTAA AGTGTCAATCA TCTGTTTACA AGGCATGCTC CCCTAGAAGT
501 TCATGCTAAA GAAGTGCCAT GGAGGCAGCA GGACAAAGTA CAGGCTAGGT
551 GGAGTCAAGC CAGGCC TAGT GCCACAGAGC AAGAGAGCAG TCTCTGACTA
601 GTAGTTAAGG GGGAAGAAAG AAAAATATTC TTCCAATTGC TTTCCAGTTC
651 TCCTTTAGGG ACAGCTTAGA ATTATTTGCA CTATTGAGTC TTCATGTTCC
701 CACTTCAAAA CAAATAGATG CTCTGAAAGC AAAC TGGCTT GAAATGGTGA
751 CACTGTCCCA CAAGCCACCA GACAATGGCA GTGTTTCAGAA CTACCTGTAT
801 ATGTATATAC CTG

Fig. 5e

Partial sequence of the non-coding RNA gene from chimpanzee

1 TTGCTGCAGA TACTACTGAC CAGACAAGCT GTTGACCAGG CACCTCCCCT
51 CCCGCCCAAA CCTTTCCCCC ATGTGGTTCGT TAGAGACAGA GCGACAGAGC
101 AGTTGAGAGG ACACTCCCGT TTTCCGGTGCC ATCAGTGCCC CGTCTACAGC
151 TCCCCCAGCT CCCCCACCT CCCCCACTCC CAACCACGTT GGGACAGGGA
201 GGTGTGAGGC AGGAGAGACA GTTGGATTCT TTAGAGAAGA TGGATATGAC
251 CAGTGGCTAT GGCCTGTGTG ATCCCACCCG TGGTGGCTCA AGTCTGGCCC
301 CACACCAGCC CCAATCCAAA ACTGGCAAGG ACGCTTCACA GGACAGGAAA
351 GTGGCACCTG TCTGCTCCAG CTCTGGCATG GCTAGGAGGG GGGAGTCCCT
401 TGAACTACTG GGTGTAGACT GGCCTGAACC ACAGGAGAGG ATGGCCCAGG
451 GTGAGGTGGC GTGGTCCATT CTCAAGGGAC GTCTTCCAAC GGGTGGCGCT
501 AGAGGCCATG GAGGCAGTAG GACAAGGCGC AGGCAGGCTG GCCCAGGGTC
551 AGGCCGGGCA GAGCACAGCG GGGTGAGAGG GATTCTTAAT CACTCAGAGC
601 AGTCTGTGAC TTAGTGAGCA GGGGAGGGGG CAAAGGGGGA GGAGAAGAAA
651 ATGTTCTTCC AGTTACTTTC CAATTCTCCT TTAGGGACAG CTTAGAATTA
701 TTTGCACTAT TGAGTCTTCA TGTTCCCACT TCAAAACAAA CAGATGCTCT
751 GAGAGCAAAC TGGCTTGAAT TGGTGACATT TAGTCCCTCA AGCCACCAGA
801 TGTGACAGTG TTGAGAACTA CCTGGATTG TATATATACC TG

Fig. 5f